

Har barnet hjertefeil?

En undersøkelse av
auskultasjonsferdighetene til
medisinstudenter og barneleger



**INSTITUTT FOR ALLMENN-
OG SAMFUNNSMEDISIN**
2008

**GEIR KRISTENSEN
HÅKON B. JEGSTAD
STUDMED V-03
UNIVERSITETET I OSLO**

Innholdsfortegnelse

Forord	3
Abstract	4
Innledning:	5
Materiale og metode	10
Resultater	18
Diskusjon	26
Referanser:	31
Vedlegg 1 : Informasjon til deltakere	32
Vedlegg 2: Svarark leger	33
Vedlegg 3: Svarark studenter	34

Forord

Denne oppgaven kom i gang i begynnelsen av mars 2006. Per Lagerløv, førsteamanuensis dr. med. Institutt for allmenn- og samfunnsmedisin ved Universitetet i Oslo, holdt et innlegg på 7. semesters prosjektdag. Dette var starten på samarbeidet som har endt i prosjektoppgaven “Har barnet hjertefeil?”

Oppgaven er basert på testing av auskultasjonsferdighetene til barneleger ved Ullevål Universitetssykehus og medisinstudenter ved Universitet i Oslo. Testmaterialet har blitt utarbeidet fra Per Lagerløvs data fra St. Hanshaugen helsestasjon.

I utgangspunktet hadde oppgaven til hensikt å finne ut av medisinstudenters auskultasjonsferdigheter. Oppgaven har vokst i størrelse og leger ved Ullevål Universitetssykehus ble inkludert i januar 2007. Det vil også bli gjort innsamling av data fra barnekardiologer ved Ullevål Universitetssykehus og ved Rikshospitalet våren 2008.

Vi ønsker å takke vår veileder Per Lagerløv for hjelp i skriveprosessen og annen veiledning, samt Leif Brunvand ved Barnesenteret Ullevål universitetssykehus for hjelp til rekruttering av deltakere og utvikling av testmateriale.

All statistisk analyse har vi gjort i Epidata/Epianalyse og Excel.

Sensor for oppgaven er seksjonsoverlege Leif Brunvand ved Barnesenteret Ullevål universitetssykehus.

Abstract

Introduction:

During routine examination of children heart murmurs are frequently encountered in primary care. It may be difficult to judge whether the murmur is innocent or due to a less serious heart disease in healthy children. Thus these children are frequently referred for further examinations. On average only ten percent of referred children are identified to have a heart defect. This low precision in discerning innocent and pathological murmurs in primary care goes against a high precision among experienced pediatricians.

Aim: To examine the precision in assessing heart sounds among medical students and doctors with varying levels of experience using sound samples reflecting the prevalence of heart murmurs in primary care.

Method and material:

From a sample of heart sounds recorded from healthy children six weeks to two years of age, two test sets with 25 sounds were constructed with different prevalence of innocent murmurs and 12 percent pathological murmurs. These sounds, with low to moderate strength, were presented to medical students and doctors to decide upon whether they were normal, innocent or pathological.

Results:

24 doctors and 35 medical students participated in the study. The positive predictive values for identifying pathological murmurs were 0.3 for both groups. The positive predictive values were lower in the test set with a high compared to a low prevalence of innocent murmurs. The doctors were somewhat better than the medical students to identify pathological murmurs, missing 22 percent compared to 38 percent. Fleiss kappa was 0.32 for doctors and 0.18 for medical students. Average intra-observatory agreements were 68 and 69 percent respectively.

Conclusion:

Experience did not improve the precision in assessing heart sounds significantly in this study. Auscultation as a sole method for discerning innocent and low intensity pathological murmurs seems to be unreliable. In primary care a better tool is needed to do appropriate screening. Doctors in secondary care must proceed with more precise examinations like echocardiography to make a safe diagnosis.

Bakgrunn

Det blir hvert år henvist hundrevis av barn til hjertepoliklinikkene ved landets sykehus. med spørsmål om organisk hjertefeil. Bakgrunnen for mange av henvisningene er funn av hjertebilyd ved auskultasjon. I følge muntlig meddelelse fra seksjonsoverlege Leif Brunvand ved Barnesenteret Ullevål Universitetssykehus (UUS), ble det i perioden januar 2000 til desember 2005 henvist 1245 barn med spørsmål om hjertepatologi. Det ble i 90% av tilfellene ikke funnet organisk hjertefeil. Tilsvarende ble det ved Haukeland Universitetssykehus i 2001 henvist 220 barn til deres barnekardiologiske poliklinikk med samme problemstilling (1). Av de 220 henviste barna ved Haukeland Univeritetssykehus fikk 22 påvist organisk hjertefeil ved klinisk undersøkelse og ekkokardiografi. Ved Universtetssykehuset i Nord-Norge hadde 99 av 106 henviste barn til hjertepoliklinikken, over en to-årsperiode, fysiologisk bilyd (2). Dette vil si at kun en liten prosentandel av barn i Norge som henvises for undersøkelse hos hjertespesialist faktisk har hjertepatologi.

Forekomsten av organisk hjertefeil varierer noe i litteraturen, men ligger mellom 0,6% (3) til over 1% (4), dersom man ekskluderer bikuspid aortaklaff og åpentstående ductus arteriosus (PDA) hos premature. I et epidemiologisk materiale fra Vestfold i perioden 1985-2005 er forekomsten av hjertefeil på 1,16%. I dette materialet, som omfattet 57000 barn, utgjør atriaseptumdefekt (ASD) og ventrikkelseptumdefekt (VSD) tilsammen 2/3 av de organiske hjertefeilene (5). Forekomsten av fysiologisk bilyd er høyere enn patologisk bilyd. I litteraturen er imidlertid tallene varierende, fra en kumulativ prevalens på 30 %, angitt i Nelson's Textbook of Pediatrics (ref 6), til opptil 80% i tidlig barndom, angitt i Practical Pediatrics (7). Punktprevalensen er lavere, men få tall finnes. I et upublisert materiale fra Lagerløv finnes en punktprevalens på 8.6% i aldersgruppen 6 uker til 2 år. ASD og VSD har samme punktum maximum som den vanligst forekommende fysiologiske bilyden, Stills (8). Dette kan gjøre vanskelig for den undersøkende lege å avgjøre om lyden er uttrykk for sykdom.

Nåværende kunnskap

Studier viser at barnekardiologer med høy sensitivitet og spesifisitet kan identifisere patologisk bilyd ved hjelp av auskultasjon (2) og klinisk undersøkelse (1) alene. Dette er i mindre grad tilfellet for andre leger. Studier i USA viser at leger under spesialisering i pediatri har suboptimale auskultasjonsevner. (9,10). Dette gjelder også indremedisinere og allmennpraktikere, både ferdige spesialister og leger under spesialisering i de samme disiplinene.

Andre studier viser at erfarne pediatere med relativt god presisjon kan skille fysiologisk fra patologisk bilyd (11,12). Hvilke kriterier som brukes på et slikt skille er ikke alltid eksplisitt forklart. Gjennomgående for de studiene vi har funnet er at styrken på bilydene som det testes på ikke er oppgitt. Fysiologiske bilyder har svak styrke (klassifisert som styrke 1-2)(8), mens patologiske bilyder også kan ha høy styrke (grad 3 eller mer). Derfor kan styrken alene være et viktig kriterium for å skille fysiologisk fra patologisk bilyd. Forekomsten av patologiske bilyder er høy i testmaterialet for disse studiene. Den positive prediktive verdi kan derfor bli betydelig høyere enn hvis testmaterialet inneholder patologiske lyder med lavere hyppighet svarende til forekommer i en populasjon av henviste barn. Prevalensen av fysiologiske bilyder vil sannsynligvis påvirke den positive prediktive verdi for påvisning av patologisk bilyd.

Det ser altså ut til at erfaring innen kardiologi er den enkeltfaktoren som har mest å si for evnen til å klassifisere hjertelyder. Studier viser da også at repetitiv mengdetrening er den beste måten å øke ferdighetsnivået på (13). Man kan derfor ikke se bort fra at mønstergjenkjennelse, pattern recognition, er viktig for å skille fysiologisk fra patologisk bilyd. I slike tilfeller benyttes ikke en analytisk tilnærming der man bevisst tar stilling til lydens plassering i syklus og lydens karaktertrekk, men legger større vekt på det intuitive inntrykk lyden gir sett i forhold til tidligere erfaring.

Auskultasjonsevnene til ikke-spesialister er ikke så dårlige som henvisningstallene skulle tilsi. En studie viser at treffsikkerheten til sisteårs medisinstudenter er på 59.1% (± 3.8), i forhold til hjertespesialister på 71.8% (± 3.5) (14). Hvorfor er det da så stor diskrepans mellom antallet henvisninger med mistanke om hjertepatologi og organisk hjertefeil hos de henviste? Grunnene kan være mange: 1) Det kan være lavere dyktighet blant norske leger enn det man finner i internasjonal litteratur. 2) Bekymring fra foreldre vil kunne være en annen viktig faktor (15). Ønsket om å avklare spørsmålet en gang for alle med ekkokardiografi, kan i så måte berolige både foreldre og legen. 3) Man kan også tenke seg at studiene som kartlegger disse evnene ikke i god nok grad representerer virkeligheten legene møter i sin praksis.

Ved de fleste barnehjertepoliklinikker i Norge gjøres ekkokardiografisk undersøkelse av alle barn som henvises med spørsmål om bilyd. Dette er uavhengig av spesialistens funn ved klinisk undersøkelse (16).

I en gjennomsnittlig norsk allmennpraksis og på helsestasjonene vil de fleste barna være hjertefriske. De aller fleste barna med alvorlig patologi får dette påvist på sykehuset. Disse barna

har enten symptomer som fører til videre undersøkelser, eller de får påvist sykdom ved standardundersøkelse på barselpostene før avreise fra sykehus. Allmennpraktikere og helsestasjonsleger møter hovedsakelig barn med en normal vekstkurve, som spiser og utvikler seg normalt. Hvis legen hos disse pasientene hører en bilyd, må legen må ta en avgjørelse om dette er noe som bør forventes eller noe som bør henvises.

Det er blitt foreslått som metode å overføre lydspor elektronisk fra allmennlege eller helsestasjonslege til barneavdeling, hvor det er av usikre auskultasjonsfunn. (2) En slik metode vil redusere antall henvisninger og føre til reduserte kostnader. For at denne teknologien skal kunne tas i bruk er man avhengig av at mottakende lege er langt dyktigere enn henvisende lege til å bedømme hjertelyder ved auskultasjon alene. Dahl et al publiserte i 2002 en studie (2) hvor barnekardiologer lyttet til opptak av hjertelyder. Denne studien konkluderte med at barnekardiologer med en høy grad av sikkerhet kunne skille fysiologisk fra patologisk bilyd. De fant en sensitivitet på 89.7% og spesifisitet på 98.2%, som vil støtte en telemedisinsk metode for å vurdere en hørt bilyd. Imidlertid har de tatt utgangspunkt i en høy (40%) forekomst av patologiske bilyder og en sammensetning av organiske hjertefeil som ikke er representativ for det som er vanlig i allmennpraksis. I deres studie ble det ikke opplyst hvilken styrke de fysiologiske og de patologiske bilydene hadde. Da det ikke ble testet spesifikt på de lydene som ville vært aktuelle for telemedisinsk vurdering, dvs friske barn med svak bilyd, vil det være vanskelig å trekke konklusjoner om denne studien gjenspeiler den virkelighet legen ville oppleve i sitt arbeid, både fra henvisende og mottakende side. Det vil derfor være interessant å ta utgangspunkt i en lydverden som likner mer på de lyder en allmennpraktiker eller helsestasjonslege møter.

I vår studie skal vi bruke et testmateriale av lyder som forsøker å gjenspeile de hjertelydene som allmennpraktikere møter ved undersøkelse av barn i egen praksis eller på helsestasjonen.

Hva er det vi ønsker å finne ut av

Vi ønsker å kartlegge auskultasjonsferdighetene til norske leger. Legene er delt inn i tre grupper: En gruppe som vil være representative for den store mengden leger, ikke-spesialister, som henviser til 2.linjetjenesten. En annen gruppe representerer leger med spesiell interesse for barnemedisin, men som ennå ikke er ferdige spesialister. Den tredje gruppen er leger med spesiell erfaring i barnekardiologi eller med lang erfaring som barnelege.

Gruppe 1: Auskultasjonsferigheter hos ikke spesialister er, så vidt vi har kunnet bringe på det rene, ikke tidligere undersøkt i Norge. Studier har vist at auskultasjonsevner til medisinstudenter sent i studiet er jevn gode med allmennpraktikere og assistentleger i pediatri og indremedisin (Ref.7).

Derfor bør medisinstudenter være en representativ gruppe for leger i 1. linjetjenesten, som er de som i hovedsak henviser til barnekardiologisk poliklinikk. Medisinstudenter er en gruppe som er lettere tilgjengelig enn allmennpraktikere som har en travel hverdag. Studentene som gruppe er plassert tettere geografisk og har en hverdag som i mindre grad er presset for tid.

Medisinerstudenter sent i studiet representerer også en mer homogen gruppe enn allmennpraktikere som kan ha vidt forskjellig erfaring med å undersøke barn avhengig av sin praksisprofil. Derfor har vi brukt 5. og 6. års medisinstudenter fra Universitet i Oslo som representanter for denne gruppen.

Gruppe 2: Barnelegene er inkludert i studien for å se om ferdighetsnivået på hjerтеаuskultasjon påvirkes av erfaring (14). Det vil være naturlig å tenke at dyktigheten vil stige med økende erfaring selv om publisert materiale ikke alltid kan vise en slik sammenheng. (11,12,14). Gruppen består av barneleger med inntil 7 års erfaring.

Gruppe 3: I denne gruppen er det barneleger med mer enn syv års erfaring. Erfaringslengden tilsvarer turnustjeneste og normert tidslengde på spesialistutdannelse i pediatri. Både i gruppe 2 og gruppe 3 er legene ansatt ved Barnesenteret Ullevål Universitetssykehus.

Metodevalg

Vi har valgt å fokusere på auskultasjonsferdigheter alene, Auskultasjonsferdigheter vil, i motsetning til full klinisk undersøkelse, være den egenskapen som vil være avgjørende ved bruk av telemedisin. Dette er også betydelig lettere å gjennomføre. Vi ønsker også å se i hvilken grad hyppigheten av fysiologisk bilyd påvirker for resultatet. Dette fordi fysiologisk bilyd er en av de viktigste faktorene som forstyrrer bildet for den undersøkende lege ved hørt bilyd. Vi har derfor komponert et testsett med hjertelyder som vi mener gjenspeiler virkeligheten i allmennpraksis. Det ble laget to seanser med henholdsvis 28% og 12% fysiologisk bilyd. Seansen med 28% forekomst av fysiologiske bilyder representerer en forekomst av fysiologisk bilyd som samsvarer med litteraturen (6). Seansen med 12% forekomst av fysiologisk bilyd er nærmere den forekomst som er funnet i upubliserte data fra Lagerløv og Brunvand. De fleste barn har ikke bilyd på hjertet. Derfor er det en høy andel av normale hjertelyder i testmaterialet. De utvalgte patologiske lydene er svake (grad 2 og 3) og har samme lokalisasjon som vanligste fysiologiske bilyd.

Hypotese

Hypotesen er at auskultasjonsferdighetene skal øke med erfaring også når man anvender testlyder som avbilder den henvisende leges hverdag. Hvis ferdighetene er betydelig høyere hos legene med lang barnemedisinsk erfaring vil disse kunne overprøve vurderingene til legene med mindre erfaring ved auskultasjon alene. Hvis deres ferdigheter ikke er tydelig bedre vil man kunne trekke i tvil om telemedisinsk metodikk kan anvendes. Lav auskultasjonsferdighet for gruppe 1 leger kan betinge at bilyder som skyldes patologi blir oversett. I slike tilfelle må man utvikle andre metoder for screening. Lav auskultasjonsferdighet for gruppe 1 leger kan også forklare hvorfor denne gruppe leger har henvisninger med mange falske positive funn. Hvis auskultasjonsferdighetene for gruppe 3 leger ikke ligger på et betydelig høyere nivå enn for gruppe 1 leger vil det forsvare at man anvender ekko-doppler metodikk rutinemessig.

Denne studien har til hensikt å teste tre grupper av leger med forskjellig erfaring i barnemedisin for deres evne til å identifisere hjertelyder hos barn som normale, fysiologiske eller patologiske i et testbatteri av hjertelyder hentet fra barn som besøkte helsestasjonen.

Materiale og metode

Lydinnsamling

For å teste deltakernes auskultasjonsevner var det behov for å utvikle et testinstrument der de utvalgte lydene skulle avspeile den forekomst av lyder man møter på en helsestasjon. Utgangspunktet var et materiale samlet inn ved St. Hanshaugen Helsestasjon i 2006. Det ble gjort opptak av hjertelydene til barn i alderen 6 uker til 2 år i forbindelse med rutinekontroller. Lydene ble tatt opp med et Littmann 4000 elektronisk stetoskop i utvidet frekvensmodus (20-1000 Hz). Dette produserer ukomprimerte lydfiler i wav-format, med 8 sekunders varighet, som så kan overføres til datamaskin for nærmere undersøkelse og analyse.

Totalt ble hjertelydene til 462 barn, av totalt 1210 barn i denne aldergruppen og dette distriktet, tatt opp. 33 av lydene ble senere fjernet på grunn av lav kvalitet eller mye støy i opptaket. Av de 429 resterende var 212 gutter og 217 jenter. 37 av disse hadde fysiologisk bilyd (Stills type). Totalt 11 barn hadde organisk hjertefeil. Av disse var det 4 som ikke hadde bilyd. 7 pasienter hadde altså organisk hjertefeil med patologisk bilyd. Av barna med patologisk bilyd hadde 4 ventrikkelseptumdefekt (VSD), én atriaseptumdefekt (ASD), én pulmonal stenose og én hypoplastisk venstre hjertesyndrom.

Utvelgelse

Det ble fra råmaterialet over valgt ut et sett lyder basert på følgende kriterier (Tabell 1): De normale og fysiologiske hjertelydene skulle ha god lyd kvalitet, med omtrent lik fordeling med hensyn på kjønn og aldersklasser. De patologiske bilydene skulle ha god lyd kvalitet, lyd styrke omtrent på nivå med fysiologisk bilyd (grad 1-2) og med punctum maximum langs venstre sternalrand. Barna med patologisk bilyd skulle også være klinisk friske med normal vekstkurve.

Kriterier	
<i>Normale/fysiologiske</i>	<i>Patologiske</i>
1) God lyd kvalitet	1) God lyd kvalitet
2) Omtrent lik fordeling mhp kjønn	2) Klinisk friske barn med normal vekstkurve
3) Omtrent likt utvalg fra alle aldersklasser	3) Lyd styrke omtrent på nivå med fysiologisk bilyd (grad 1-2), men har ett unntak (grad 3)
	4) Lyd med punctum maximum langs venstre sternalrand

Tabell 1: Kriterier for utvelgelse av hjertelyder til testmaterialet i studien .

Vi ønsket de hyppigst forekommende hjertefeilene representert i testmaterialet, i form av VSD og ASD. Disse har en prevalens på hhv 56% og 10 % av organiske hjertefeil i følge en epidemiologisk studie fra Vestfold (4). Disse har også et punktum maximum som har samme lokalisasjon som Stills fysiologiske bilyd.

Lydene ble klassifisert uavhengig av 2 pediatere, hvorav 1 barnekardiolog og 1 helsestasjonslege, basert på lydopptakene. Deres felles bedømming ble testens gullstandard. De hadde tilgang på begrenset klinisk informasjon. Lydene ble lyttet til i Littmann ® Sound Analysis Software v2.0c utviklet av Stethographics Inc. Man har der muligheten til å velge mellom ulike frekvensområder, samt visning av fonokardiogram og spektrogram. I tillegg kunne avspillingshastigheten justeres. For de patologiske lydene er det også gjort ekkokardiografisk undersøkelse.

Vi endte opp med 31 ulike lyder, hvorav 3 patologiske, 7 fysiologiske og 21 normale (Tabell 2). Fra disse 31 lydene ble det satt sammen to lydsett, kalt henholdsvis seanse A og B. I seanse A var det 15 normale, 7 fysiologiske og 3 patologiske. I seanse B var det 19 normale, 3 fysiologiske og 3 patologiske. I seanse B var alle de fysiologiske og patologiske bilydene identiske med de i seanse A. Av de 21 normale lydene er 13 felles i de to seansene. Totalt sett endte vi opp med 19 par identiske lyder og 6 unike lyder i hver seanse, totalt 25 lyder i hver seanse.

Sammensetning av testmaterialet

<i>Hjertelyd</i>	<i>Lydbeskrivelse</i>
Normale	6 uker: 6 jenter 6 mnd: 2 jenter, 2 gutter 1 år: 1 jente, 2 gutter 2 år: 3 jenter, 5 gutter
Fysiologiske	6 uker: 1 gutt (Stills grad I) og 1 jente (Stills grad I) 6 mnd: 1 gutt (grad I) og 2 jenter (Stills grad II) 2 år: 2 gutter (Stills grad I og II)
Patologiske	6 uker: gutt, VSD grad II-III 6 mnd: Jente, ASD grad III 1 år: Jente, VSD, grad II

Tabell 2: Viser alders- og kjønnsfordeling, samt lydens type og styrke på barna som lyden er tatt opp på.

Begrunnelsen for å velge totalt 50 lyder var i hovedsak praktiske hensyn. Dels på grunn av hensyn til tidsforbruk, dels for å ikke slite ut deltakerne med et stort antall lyder. Dette medførte også at prevalensen av patologiske lyder ble uvanlig høy i vårt testmateriale, rundt 10 % mot vanligvis 1 % forekomst av organisk hjertefeil. Et visst antall patologiske bilyder var nødvendig for å kunne si noe om evnen til å skille de patologiske fra de fysiologiske bilydene.

Ved å dele materialet i 2 seanser fikk vi muligheten til å sammenlikne presisjonen når forekomsten av fysiologiske bilyder varierer. Forekomsten av fysiologiske bilyder i seanse A var på 28%, tilsvarende det man finner i litteraturen (6). I seanse B har vi en forekomst på 12%, nærmere det man finner i foreløpig upubliserte data fra Lagerløv og Brunvand.

Ved å gjenta flere av hjertelydene i de to seansene fikk vi også muligheten til å vurdere reliabiliteten til deltakerne ved å sjekke om like lyder ble vurdert likt av samme deltaker.

Lydene i de to seansene ble randomisert med tanke på tidsrekkefølge.

Rekruttering/Deltakerne

Deltakerne i studien ble rekruttert fra medisinstudiet i Oslo og fra Barnesenteret ved Ullevål Universitetssykehus. Kravet til studentene var at pediatriundervisningen var kommet godt i gang. Det vil si at deltakerne var kommet i gang med 9. semester av studiet. I tillegg ble det rekruttert deltakere som hadde kommet lengre i studiet, og da hadde fullført sin pediatriundervisning. Rekruttering av studentene foregikk ved utsendelse av e-post, oppmøte på forelesninger og plakater ved studentområdet på Rikshospitalet. På 9. semester ved UiO er det også utvekslingsstudenter fra en rekke land. Disse har blitt ekskludert fra

studien på grunn av betydelig variasjon i klinisk erfaring. Totalt ble det sendt ut invitasjon om deltakelse til 387 studenter.

Studentene ved UiO har etter innføringen av Oslo 96-reformen fått klinisk undervisning tidlig i studiet. Deltakerne har vært gjennom alle spesialistdisipliner innen medisinen.

For legene var kravet at de var ansatt ved Barnesenteret ved UUS. De ble rekruttert via e-post, oppmøte på internundervisning og morgenmøter. Vi kontaktet også en seksjonsoverlege ved barnesenteret som bidro i rekrutteringsprosessen. Det er ca 70 leger ansatt ved barnesenteret på UUS.

Informasjon til deltakerne

Følgende informasjon ble gitt før lyttingen begynte (Tabell 3): Det er 2 seanser som inneholder 25 lydspor hver. De 2 seansene har forskjellig forekomst av bilyder. Lydene er tatt opp på en helsestasjon i Oslo. Alle bilyder, fysiologiske og patologiske, var funnet som et tilfeldig funn ved den vanlige helseundersøkelsen. Barna var klinisk friske. D.v.s. at det ingen av barna hadde "failure to thrive", redusert aktivitet, ernæringsvansker etc.

Informasjon til deltakerne
<ul style="list-style-type: none">- Det er 2 seanser som inneholder 25 lydspor hver.- De 2 seansene har forskjellig forekomst av bilyder- Lydene er tatt opp på en helsestasjon i Oslo- Alle bilyder, fysiologiske og patologiske, var funnet som et tilfeldig funn ved den vanlige helseundersøkelsen.- Barna var klinisk friske. D.v.s. at det ingen av barna hadde "failure to thrive", redusert aktivitet, ernæringsvansker etc.

Tabell 3: Informasjonen som ble gitt til alle deltakerne før lyttingen startet.

Svarskjemaet som ble delt ut til studentene inneholdt spørsmål om kull, kjønn og alder (Vedlegg 1). Svarskjemaet til legene inneholdt spørsmål om antall år siden medisinsk embetseksamen, type stilling og kjønn. Hoveddelen, som var lik for begge grupper, inneholdt 3 forskjellige svaralternativer: Én kolonne for normal bedømmelse av lyden, én kolonne for fysiologisk bedømmelse og én kolonne for patologisk bedømmelse. Deltagerne ble bedt om å skrive inn nummeret på lyden de hørte på i en av de 3 kategoriene.

Utstyr

Programvaren som ble brukt av deltakerne var iTunes versjon 7.0 og 7.1, installert på bærbare datamaskiner. iTunes er et lett tilgjengelig lydsporavspillingsprogram som kan spille lydspor av typen wav-filer, som vi bruker i dette forsøket. Deltagerne hadde ikke mulighet å regulere hastigheten på lydavspillingen. Opptaket av lydene ble gjort i utvidet frekvensmodus, 20-1000 Hz. Det var ikke mulig å regulere frekvensområdet for deltakerne. Det eneste visuelle hjelpemiddelet deltakerne hadde var et bilde av lydsyklusens varighet, som var på 8 sekunder. Denne forløp som en linje øverst på iTunes-vinduet. Programmet var satt opp slik at lydsporene spilte repetitivt. For å komme til neste lyd måtte man klikke på den neste lyden. Underveis i lyttingen kunne deltakerne justere volumet på lydene. Underveis sto man fritt til å velge rekkefølgen på lydene, men det ble oppfordret til å lytte til lydene kronologisk.

De innebygde høyttalerne på datamaskinen var ikke gode nok for å kunne spille av lydene med tilstrekkelig styrke og frekvensområde, spesielt bassområdet. For å sikre at lydene ble tilfredsstillende reproduisert, ble det derfor benyttet tett omsluttende hodetelefoner.

Før deltakerne begynte lyttingen måtte de lytte til 3 referansetoner. Disse var på 30 Hz, 150 Hz og 800 Hz. Bakgrunnen for dette var å forsikre oss om at de hadde normal hørsel innenfor frekvensområdet som lydene var tatt opp i.

Gjennomføring og tilbakemelding

Gjennomføringen av lyttingen ble for studentene gjort på et grupperom på Rikshospitalet, for legene ble lyttingen gjort på den enkelte leges kontor. Det ble for begge grupper passet på at det var stille og at det ikke var forstyrrende elementer tilstede under gjennomføringen. Det ble ikke lagt noen tidsbegrensing på lyttingen, altså kunne deltakerne bruke så kort eller lang tid de følte var nødvendig.

Tilbakemelding ble gitt umiddelbart etter besvarelsen til den enkelte deltaker. Den ble gitt i form av en prosentvis treffsikkerhet og på hvor de lydene som ikke var klassifisert rett var plassert. Noen deltakere ønsket å lytte på nytt på lyder de hadde vært usikre på. Dette ble det gitt tillatelse til.

Statistisk metode:

Dataene ble ført inn i EpiData v.1 og analyse ble hovedsakelig gjort i søsterprogrammet EpiAnalyse v.1. Bakgrunnen for valget av EpiData var at det er gratis tilgjengelig på nett, noe som åpner for datainnsamling og -analyse uten å ha tilgang på dyre programlisenser. Det brukes bl.a. av WHO i deres epidemiologiske studier og mottar støtte bl.a. fra Dnlf. EpiData er kompatibelt med mer avanserte statistikkprogrammer som SPSS.

Vi ønsket å sammenlikne resultatene til legene og studentene, samt resultatene i seanse A og seanse B. For å vurdere i hvilken grad erfaring har noe å si for auskultasjonsferdighetene har vi også delt opp legematerialet i to mindre deler, avhengig av erfaringsnivå, nærmere bestemt mer eller mindre enn 7 års erfaring. Valget av 7 år som grense skyldes dels at dette er medianen i materialet, noe som gjør at vi får to like store grupper. I tillegg svarer 7 års erfaring lengden på turnustjeneste og 5 års spesialistutdanning i pediatri.

Det ble regnet ut spesifisitet og sensitivitet for henholdsvis normal, fysiologisk og patologisk lyd, for å undersøke hvor dyktige man er på å identifisere hver av de tre typene lyder. Disse målene er uavhengige av prevalens og bør derfor ikke forandres i større grad i de to seansene. Sensitivitet og spesifisitet er parametre som kan regnes ut fra 2x2-tabeller der deltakernes besvarelser vurderes i forhold til en gullstandard (Figur 1).

		Gullstandard	
		Sann	Positiv
Test-utfall	Positiv	Sann positiv	Falsk Positiv
	Negativ	Falsk Negativ	Sann negativ

Figur 1: 2x2-tabell. Grunnlag for beregning av sensitivitet, spesifisitet og prediktive verdier.

Sensitivitet er et mål på i hvilken grad en test korrekt identifiserer de positive tilfellene. I vårt tilfelle hvor stor andel av de med henholdsvis normal, fysiologisk og patologisk lyd, man klarer å identifisere.

Spesifisitet er et mål på i hvilken grad en test korrekt identifiserer de negative tilfellene.

I tillegg har vi regnet positiv prediktiv verdi (PPV) for identifisering av patologiske bilyder. Den uttrykker den andelen av de positive testresultatene som faktisk er syke. Positiv prediktiv verdi vil her indikere hvor mange av de man henviser med mistanke om organisk hjertefeil, som faktisk har det.

Prediktive verdier er i motsetning til sensitivitet og spesifisitet relatert til prevalensen, slik at PPV blir høyere med økende prevalens. Man ville derfor forvente en økning av PPV når antallet fysiologiske bilyder i forhold til de patologiske faller i seanse B, fordi antallet forvirrende elementer (fysiologiske bilyder) blir redusert.

For både sensitivitet, spesifisitet og positiv prediktiv verdi gjelder det at besvarelsene testes opp mot en gullstandard, som selvsagt kan være mer eller mindre pålitelig.

For å undersøke i hvilken grad eventuelle forskjeller er reelle, *signifikans*, har vi benyttet student t-test, som så har blitt omregnet til p-verdier. Grensen for signifikans er satt til $p=0,05$.

I figur 2 sees formlene for de ulike statistiske beregningene. I formelen for student's t-test står \bar{X}_1 og \bar{X}_2 for gjennomsnittene i hver gruppe som sammenliknes. s^2_1 og s^2_2 er variansen i de to gruppene. n_1 og n_2 er størrelsen på hver av gruppene.

For å teste interobservatørreliabilitet har vi brukt Fleiss kappa-test. Denne kan brukes til å teste graden av samsvar på besvarelsene til et større antall deltakere, i motsetning til Cohens kappa, som tester to og to testere (17) Kappa-testing vurderer hvor likt man som gruppe vurderer en lyd, utover det man ville forvente ved tilfeldighet. Det blir derfor et mål på samsvar innen for gruppen som testes. Testen tar ikke hensyn til hva som er riktig i forhold till gullstandard, kun hvor likt man besvarer.

Kappaverdier kan tolkes etter retningslinjene til Landis & Koch (Figur 2) (18).

<i>Value of κ</i>	<i>Strength of agreement</i>
<0	Poor
0-0,20	Slight
0,21-0,40	Fair
0,41-0,60	Moderate
0,61-0,80	Substantial
0,81-1,00	Almost perfect agreement

Figur 2: Tolkning av κ -verdier ved mål av interobservatør-reliabilitet.

Formelen for beregning av kappa står også i figur 3. Her er p_0 observert samsvar, p_e er samsvar ved tilfeldig besvarelse. $1-p_e$ er maksimalt samsvar utover tilfeldig besvarelse. For statistiske beregninger (Figur 2)

$$\text{sensitivitet} = \frac{\text{antall sann positiv}}{\text{antall sann positiv} + \text{antall falsk negativ}}$$

$$\text{spesifisitet} = \frac{\text{antall sann negativ}}{\text{antall sann negativ} + \text{antall falsk positiv}}$$

$$\text{PPV} = \frac{\text{antall sann positiv}}{\text{antall sann positiv} + \text{antall falsk positiv}}$$

$$\text{Student's } t\text{-test: } t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\text{Kappa-test: } \kappa = \frac{p_0 - p_e}{1 - p_e}$$

Figur 3: Statistiske beregninger. Formelene for utregning av de statistiske verdiene som er brukt i studien.

I seanse B er de siste 11 lydene normale. Man kan tenke seg at dette kunne føre til et fall i treffsikkerheten på slutten, dels på grunn av lyttetretthet mot slutten av testen og dels at deltakerne ville forvente at en bilyd ville dukke opp etter mange normale hjertelyder. For å vurdere om lydenes plassering i testmaterialet har noe å si for vurderingen, har vi derfor sett på om samsvaret mellom de identiske lydene i seanse A og B endrer seg i løpet av lyttingen. Dersom samsvaret holder seg stabilt gjennom seanse B, vil det styrke reliabiliteten i besvarelsene.

Resultater

Deltakerne

Det deltok 24 leger (Tabell 5). Av disse var 11 menn (45,8%) og 13 kvinner (54,2%). 3 av legene jobbet på hjerteseksjonen, 2 var i overlegestilling og 1 var i fordypningsstilling. 21 leger jobbet ikke i hjerteseksjonen. 8 av deltakerne er spesialister i pediatri, 16 er ikke ferdige spesialister. 6 av legene som deltok jobbet i overlegestilling, 17 jobbet som assistentleger, hvorav 4 var i fordypningsstilling. 1 lege svarte ikke på spørsmål om stilling. På grunn av stor spredning på yrkeslengde er median et bedre mål på erfaring enn gjennomsnitt. Median arbeidslengde var på gjennomføringstidspunktet 7 år.

Totalt deltok det 35 medisinstudenter (Tabell 6). Det var 15 menn (42,9%) og 18 (51,4%) kvinner. 2 (5,7%) unnlot å svare på spørsmål om kjønn. På gjennomføringstidspunktet var 11 av studentene på 9. semester, 15 studenter på 10. semester, 6 studenter på 11. semester og 2 studenter på 12. semester. 1 deltager unnlot å svare. Gjennomsnittlig alder på studentene var 27,5 år på gjennomføringstidspunktet.

Gjennomsnittlig gjennomføringstid var ca. 30 minutter for legene og ca. 45 min. for studentene.

Deltakerdetaljer, leger.

Stilling	Antall	Kjønn (Menn/Kvinner)		Spesialister	Gjennomsnittlig erfaring i år (min.-maks.)
Overlege	6	5	1	6	22 (14,6-29,4)
Ass. Lege	17	6	12	2	7,3 (5,1-9,5)
Ubesvart	1		1	0	
Total	24	11	12	8	11.0 (7,5-14,5)

Tabell 4: Oversikt over de deltakende legene.

Deltakerdetaljer, studenter					
Semester	Antall deltakere	Kjønn *		Potensielle deltakere	% deltagelse
		(Menn/Kvinner)			
9	11	5	5	84	13,1
10	15	5	10	101	14,9
11	6	3	3	108	5,6
12	2	2	0	94	2,1
Total	35	15	18	387	9,0

Tabell 5: Oversikt over de deltagende studentene. * 2 unnlot å svare på spørsmål om kjønn.

Alle deltakerne, leger og studenter, hørte de tre referansetonene på 30, 150 og 800 Hz. Dette vil si at hørselen mest sannsynlig ikke er en faktor som påvirker resultatet.

Data, leger vs. studenter

Fordelingen av svarene for legene og studentene kan sees i henholdsvis tabell 6 og tabell 7, samt figur 4 og 5. Hver enkelt lyd er navngitt med en kode. Den første bokstaven angir i hvilken seanse lyden befant seg i (a eller b). Tallet angir den kronologiske plasseringen av lyden innenfor seansen. Den siste bokstaven angir hvordan lyden er bedømt i henhold til gullstandard (n=normal, f=fysiologisk, p=patologisk).

Besvarelser leger

Leger	N=24	Beskrivelse av hjertelyd							
Lyd	Gullstandard	Normal		Fysiologisk		Patologisk		Uavklart	Ubesvart
		n	%	n	%	n	%		
a2n	Normal	13	54,2%	8	33,3%	3	12,5%		
a3n	Normal	18	75,0%	2	8,3%	4	16,7%		
a6n	Normal	21	87,5%	2	8,3%	1	4,2%		
a7n	Normal	23	95,8%	1	4,2%	0	0,0%		
a8n	Normal	11	45,8%	10	41,7%	3	12,5%		
a9n	Normal	18	75,0%	2	8,3%	3	12,5%		1
a10n	Normal	8	33,3%	9	37,5%	7	29,2%		
a14n	Normal	19	79,2%	5	20,8%	0	0,0%		
a15n	Normal	16	66,7%	7	29,2%	1	4,2%		
a16n	Normal	12	50,0%	9	37,5%	3	12,5%		
a17n	Normal	12	50,0%	2	8,3%	10	41,7%		
a20n	Normal	22	91,7%	1	4,2%	1	4,2%		
a23n	Normal	19	79,2%	4	16,7%	1	4,2%		
a24n	Normal	0	0,0%	7	29,2%	17	70,8%		
a25n	Normal	17	70,8%	3	12,5%	4	16,7%		
b1n	Normal	16	66,7%	4	16,7%	4	16,7%		
b2n	Normal	18	75,0%	6	25,0%	0	0,0%		
b3n	Normal	11	45,8%	7	29,2%	6	25,0%		
b4n	Normal	15	62,5%	7	29,2%	2	8,3%		
b6n	Normal	3	12,5%	8	33,3%	13	54,2%		
b8n	Normal	14	58,3%	7	29,2%	3	12,5%		
b9n	Normal	18	75,0%	3	12,5%	3	12,5%		
b12n	Normal	18	75,0%	3	12,5%	2	8,3%		1
b15n	Normal	18	75,0%	5	20,8%	0	0,0%		1
b16n	Normal	18	75,0%	6	25,0%	0	0,0%		
b17n	Normal	16	66,7%	6	25,0%	2	8,3%		
b18n	Normal	21	87,5%	2	8,3%	1	4,2%		
b19n	Normal	19	79,2%	4	16,7%	1	4,2%		
b20n	Normal	16	66,7%	3	12,5%	4	16,7%		1
b21n	Normal	10	41,7%	11	45,8%	3	12,5%		
b22n	Normal	15	62,5%	5	20,8%	4	16,7%		
b23n	Normal	10	41,7%	7	29,2%	7	29,2%		
b24n	Normal	9	37,5%	11	45,8%	3	12,5%		1
b25n	Normal	5	20,8%	6	25,0%	13	54,2%		
a1f	Fysiologisk	2	8,3%	9	37,5%	13	54,2%		
a4f	Fysiologisk	1	4,2%	11	45,8%	12	50,0%		
a5f	Fysiologisk	2	8,3%	7	29,2%	15	62,5%		
a12f	Fysiologisk	0	0,0%	11	45,8%	13	54,2%		
a18f	Fysiologisk	1	4,2%	3	12,5%	20	83,3%		
a19f	Fysiologisk	2	8,3%	12	50,0%	10	41,7%		
a22f	Fysiologisk	0	0,0%	9	37,5%	15	62,5%		
b5f	Fysiologisk	1	4,2%	8	33,3%	15	62,5%		
b10f	Fysiologisk	0	0,0%	3	12,5%	21	87,5%		
b11f	Fysiologisk	0	0,0%	6	25,0%	18	75,0%		
a11p	Patologisk	4	16,7%	10	41,7%	10	41,7%		
a13p	Patologisk	1	4,2%	3	12,5%	20	83,3%		
a21p	Patologisk	0	0,0%	0	0,0%	24	100,0%		
b7p	Patologisk	1	4,2%	9	37,5%	14	58,3%		
b13p	Patologisk	0	0,0%	1	4,2%	23	95,8%		
b14p	Patologisk	1	4,2%	2	8,3%	21	87,5%		

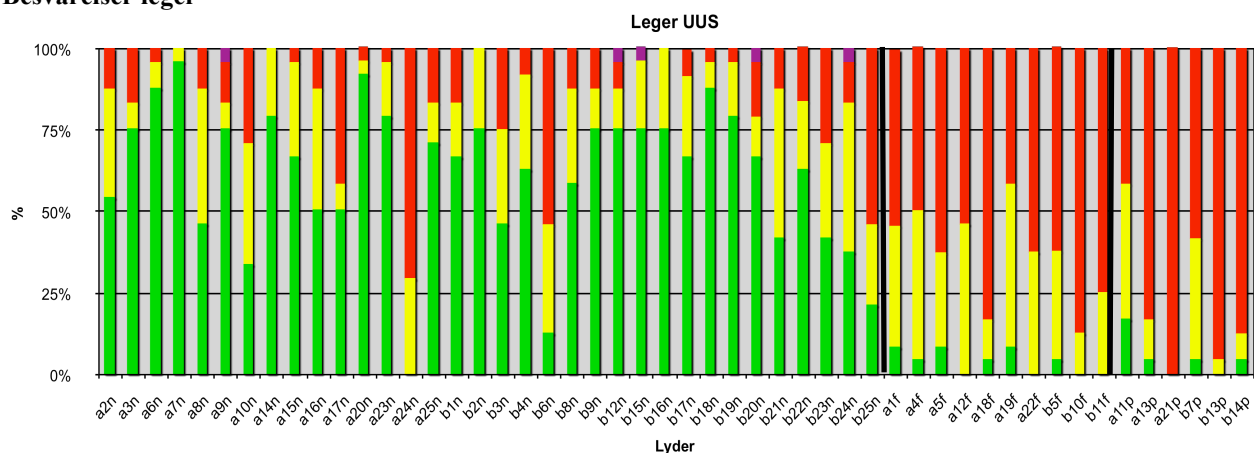
Tabell 6: Viser antall og prosentvis besvarelse av hver enkelt hjertelyd, samt vurdering bedømt etter gullstandard. Lydkoden i kolonnen lengst til venstre henviser til seanse, nr. i seansen og bedømmelse etter gullstandard. Således er f.eks. a1f plassert i seanse A som lyd nr 1 og den er bedømt til å være fysiologisk.

Besvarelser studenter

Studenter	N=35	Beskrivelse av hjertelyd							
		Normal		Fysiologisk		Patologisk		Uavklart	Ubesvart
		N	%	N	%	N	%		
a2n	Normal	23	65,7%	8	22,9%	4	11,4%		
a3n	Normal	22	62,9%	3	8,6%	10	28,6%		
a6n	Normal	31	88,6%	1	2,9%	3	8,6%		
a7n	Normal	30	85,7%	2	5,7%	3	8,6%		
a8n	Normal	26	74,3%	4	11,4%	5	14,3%		
a9n	Normal	26	74,3%	5	14,3%	4	11,4%		
a10n	Normal	19	54,3%	9	25,7%	7	20,0%		
a14n	Normal	30	85,7%	5	14,3%	0	0,0%		
a15n	Normal	26	74,3%	4	11,4%	4	11,4%	1	
a16n	Normal	20	57,1%	8	22,9%	7	20,0%		
a17n	Normal	21	60,0%	6	17,1%	8	22,9%		
a20n	Normal	28	80,0%	4	11,4%	3	8,6%		
a23n	Normal	27	77,1%	3	8,6%	5	14,3%		
a24n	Normal	7	20,0%	15	42,9%	12	34,3%		1
a25n	Normal	23	65,7%	5	14,3%	6	17,1%		1
b1n	Normal	19	54,3%	10	28,6%	6	17,1%		
b2n	Normal	25	71,4%	8	22,9%	2	5,7%		
b3n	Normal	26	74,3%	7	20,0%	2	5,7%		
b4n	Normal	20	57,1%	8	22,9%	7	20,0%		
b6n	Normal	16	45,7%	8	22,9%	11	31,4%		
b8n	Normal	19	54,3%	9	25,7%	7	20,0%		
b9n	Normal	25	71,4%	5	14,3%	5	14,3%		
b12n	Normal	23	65,7%	4	11,4%	8	22,9%		
b15n	Normal	32	91,4%	1	2,9%	2	5,7%		
b16n	Normal	28	80,0%	5	14,3%	2	5,7%		
b17n	Normal	25	71,4%	4	11,4%	6	17,1%		
b18n	Normal	27	77,1%	4	11,4%	4	11,4%		
b19n	Normal	28	80,0%	3	8,6%	4	11,4%		
b20n	Normal	27	77,1%	6	17,1%	2	5,7%		
b21n	Normal	23	65,7%	4	11,4%	8	22,9%		
b22n	Normal	20	57,1%	8	22,9%	7	20,0%		
b23n	Normal	20	57,1%	8	22,9%	6	17,1%	1	
b24n	Normal	23	65,7%	7	20,0%	5	14,3%		
b25n	Normal	11	31,4%	6	17,1%	18	51,4%		
a1f	Fysiologisk	5	14,3%	6	17,1%	24	68,6%		
a4f	Fysiologisk	9	25,7%	14	40,0%	12	34,3%		
a5f	Fysiologisk	10	28,6%	13	37,1%	12	34,3%		
a12f	Fysiologisk	9	25,7%	9	25,7%	17	48,6%		
a18f	Fysiologisk	4	11,4%	9	25,7%	22	62,9%		
a19f	Fysiologisk	10	28,6%	8	22,9%	17	48,6%		
a22f	Fysiologisk	8	22,9%	14	40,0%	13	37,1%		
b5f	Fysiologisk	2	5,7%	12	34,3%	21	60,0%		
b10f	Fysiologisk	6	17,1%	11	31,4%	18	51,4%		
b11f	Fysiologisk	10	28,6%	7	20,0%	18	51,4%		
a11p	Patologisk	7	20,0%	19	54,3%	9	25,7%		
a13p	Patologisk	2	5,7%	6	17,1%	27	77,1%		
a21p	Patologisk	2	5,7%	5	14,3%	28	80,0%		
b7p	Patologisk	12	34,3%	10	28,6%	13	37,1%		
b13p	Patologisk	2	5,7%	5	14,3%	28	80,0%		
b14p	Patologisk	5	14,3%	5	14,3%	24	68,6%	1	

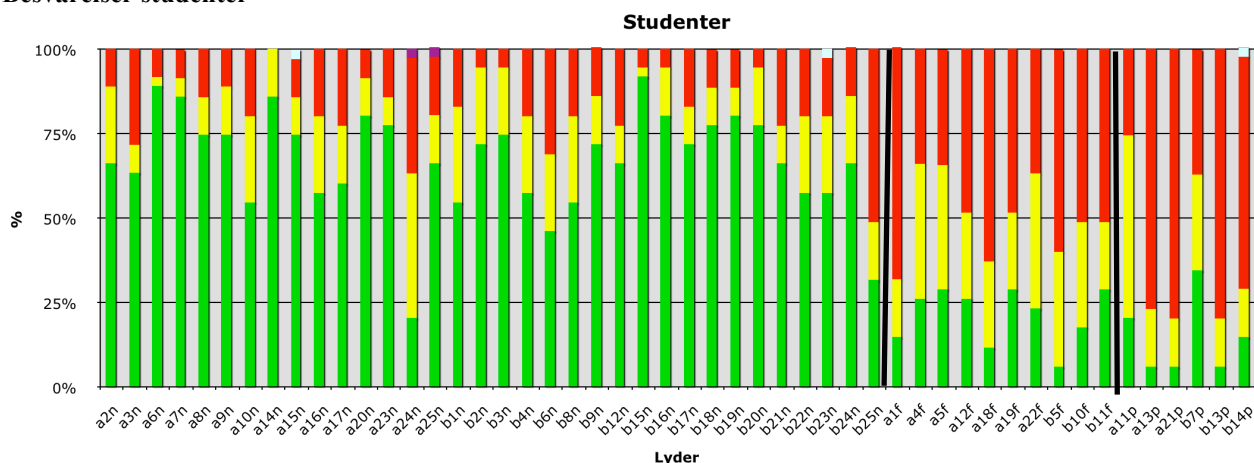
Tabell 7: Viser antall og prosentvis besvarelse av hver enkelt hjertelyd, samt vurdering bedømt etter gullstandard. Lydkoden i kolonnen lengst til venstre henviser til seanse, nr. i seansen og bedømmelse etter gullstandard. Således er f.eks. a1f plassert i seanse A som lyd nr 1 og den er bedømt til å være fysiologisk.

Besvarelser leger



Figur 4: Viser testpersonenes samlede vurdering av lydene som henholdsvis normale, fysiologiske eller patologiske. Disse er presentert i figuren med grønn, gul og rød farge, respektivt for hver enkelt lyd. For hver lyd, angitt på x-aksen, er andel av testpersonene som har vurdert den som normal (grønt), fysiologisk (gult) eller patologisk (rødt) angitt som prosent på y-aksen. Lydene i venstre felt av figuren er som standard vurdert til normale, lydene i midtre felt er vurdert som fysiologiske og lydene i høyre felt er vurdert som patologiske.

Besvarelser studenter



Figur 5: Svarfordeling studenter. Viser testpersonenes samlede vurdering av lydene som henholdsvis normale, fysiologiske eller patologiske. Disse er presentert i figuren med grønn, gul og rød farge, respektivt for hver enkelt lyd. For hver lyd, angitt på x-aksen, er andel av testpersonene som har vurdert den som normal (grønt), fysiologisk (gult) eller patologisk (rødt) angitt som prosent på y-aksen. Lydene i venstre felt av figuren er som standard vurdert til normale, lydene i midtre felt er vurdert som fysiologiske og lydene i høyre felt er vurdert som patologiske.

Figur 4 og 5 viser fordelingen av besvarelsene for hver lyd. Figur 4 viser legenes svar og figur 5 viser studentenes svar. Om alle deltakerne hadde svart rett ville det venstre feltet vært bare grønt, det midtre feltet bare gult og det høyre feltet ville vært kun farget med rødt. Y-aksen viser prosentvis fordeling av de 3 svaralternativene. X-aksen viser lydene i kronologisk rekkefølge innen hver kategori, normal, fysiologisk og patologisk.

Statistiske beregninger

Sensitivitet for legene, når begge seansene, A og B er slått sammen, er for normale hjertelyder 0,62, for fysiologiske bilyder 0,30 og for patologiske bilyder 0,78. Spesifisiteten er for normale hjertelyder 0,96, for fysiologiske bilyder 0,78 og for patologiske bilyder 0,74. Positiv prediktiv verdi er 0,32 og treffsikkerhet er 0,58.

Sensitivitet for studentene når begge seanser er slått sammen er for normale hjertelyder 0,67, for fysiologiske bilyder 0,29 og for patologiske bilyder 0,62. Spesifisiteten er for normale hjertelyder 0,82, for fysiologiske bilyder 0,82 og for patologiske bilyder 0,76. Positiv prediktiv verdi er 0,29 og treffsikkerhet er 0,59. (Tabell 8)

	Resultater leger og studenter			
	Leger (N=24)		Studenter (N=35)	
	Mean	95% CI	Mean	95% CI
Sensitivitet N	0,62	0,56-0,68	0,67	0,60-0,73
Spesifisitet N*	0,96	0,94-0,99	0,82	0,76-0,88
Sensitivitet F	0,30	0,22-0,38	0,29	0,22-0,37
Spesifisitet F	0,78	0,75-0,81	0,82	0,78-0,86
Sensitivitet P*	0,78	0,71-0,85	0,62	0,54-0,70
Spesifisitet P	0,74	0,69-0,78	0,76	0,71-0,81
Positiv prediktiv verdi	0,32	0,28-0,36	0,29	0,25-0,32
Treffsikkerhet	0,58	0,53-0,63	0,59	0,54-0,64

Tabell 8: *= signifikant. Tabellen viser gjennomsnittlig sensitivitet og spesifisitet på normale hjertelyder, fysiologiske- og patologiske bilyder. I tillegg viser tabellen positiv prediktiv verdi og treffsikkerhet.

Signifikans leger vs. studenter, sammenslåtte data seanser A og B

Det er regnet ut p-verdier ved hjelp av students t-test for sammenslåtte data mellom studenter og leger. Det eneste funnet hvor det er signifikante forskjeller mellom leger og studenter er på spesifisitet på normale (0,96 vs 0,82, $p=0,000578$) og sensitivitet på patologisk lyd (0,78 vs 0,62, $p=0,0067$).

Resultater utfra erfaring

Ved sammenlikning av leger med ulikt erfaringsnivå har vi sett på PPV og total treffsikkerhet. Treffsikkerheten steg ikke signifikant, men positiv prediktiv verdi steg fra 0,28 til 0,36, som er signifikant ($p=0,0022$). I tillegg har vi sett på spesifisitet normal hjertelyd og sensitivitet patologisk bilyd, da disse var signifikant forskjellige når legegruppen som helhet ble sammenliknet med studentene. (Tabell 9).

Resultater utfra erfaring

	Leger erfaring ≤ 7 år (3-7 år)		Leger erfaring ≥ 8 år (8-29 år)	
	N=12		N=12	
	Mean	CI 95%	Mean	CI 95%
Spesitivitet N	0,93	0,88-0,98	0,99	0,98-1,00
Sensitivitet P	0,69	0,50-0,79	0,78	0,71-0,85
PPV*	0,28	0,23-0,33	0,36	0,31-0,42
Treffsikkerhet	0,56	0,48-0,63	0,60	0,53-0,67

Tabell 9: *=signifikant forskjell. Legegruppen er her delt opp i to grupper basert på erfaring. Tabellen viser gjennomsnittlig positiv prediktiv verdi, treffsikkerhet, sensitivitet n og Spesifisitet P, med 95% konfidensintervall for de to gruppene.

Interobservatørreliabilitet

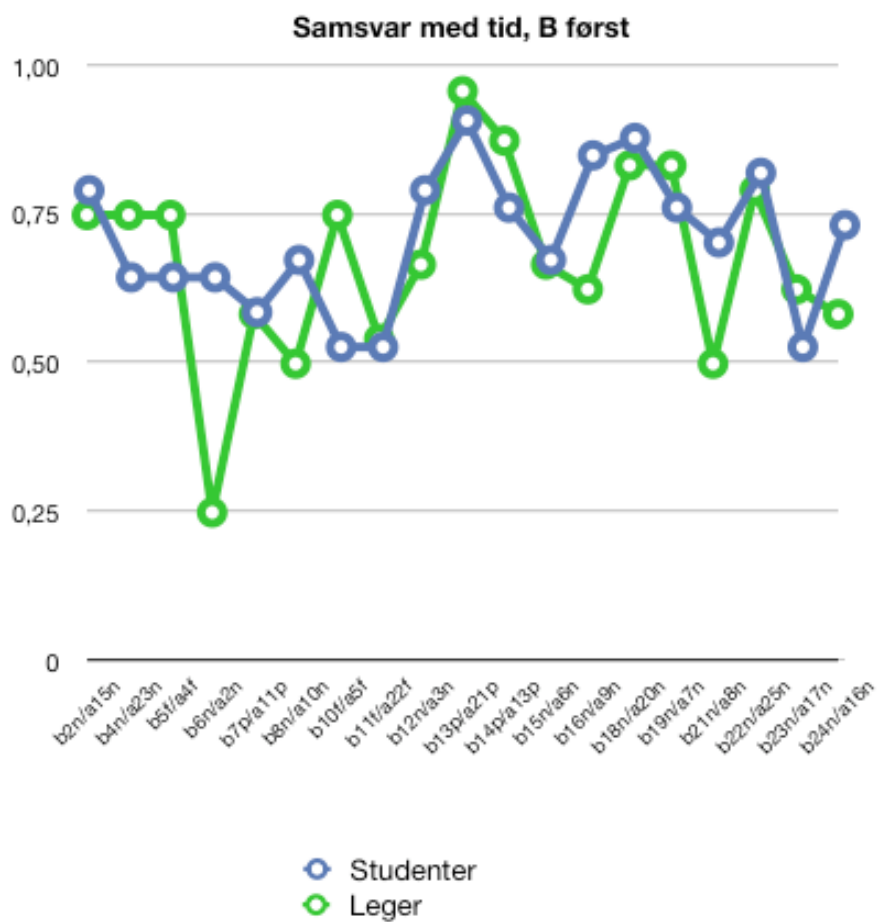
Inter-observatør kappa er 0,18 for studentene og 0,32 for legene.

Sammenlikning av seanse A og seanse B

Det er regnet ut sensitivitet, spesifisitet, positiv prediktiv verdi og treffsikkerhet for normal, fysiologisk og patologisk lyd for hver seanse separat. For studentene var det ingen signifikante forskjeller på seanse A og B, med unntak av positiv prediktiv verdi, som økte fra 0,24 i seanse A til 0,33 i seanse B ($p=0,0056$). Det samme gjaldt legene, hvor positiv prediktiv verdi økte fra 0,29 i seanse A til 0,36 i seanse B ($p=0,036$). I tillegg var det signifikant forskjell på sensitivitet fysiologisk bilyd, som sank fra 0,369 i seanse A til 0,240 i seanse B ($p=0,046$).

Samsvar mellom A og B

Gjennomsnittlig enighet mellom like lyder i seanse A og B for studentene var 69%, spredning (CI-95 52,6%-89,5%). Legene hadde en gjennomsnittlig enighet på 67,5% (CI-95% 47,4%-84,2%). Figur 6 gir en visuell fremstilling av intraobservatør-reliabiliten, og hvordan denne endrer seg i løpet av seanse B. For å vurdere om lydets plassering i seansen har hatt betydning for besvarelsen har vi plassert lydene i seanse B som er identiske med seanse A, i kronologisk rekkefølge langs x-aksen slik at denne representerer tidsforløpet i seanse B. Samsvar med tilsvarende lyder i seanse A kan leses på Y-aksen.



Figur 6: Visuelt mål på intraobservatørreliabilitet. X-aksen viser lydene i seanse B i kronologisk rekkefølge, slik at denne blir en tidsakse. Y-aksen viser gjennomsnittelig samsvar med tilsvarende lyder i seanse A.

Diskusjon

Våre funn

I vår studie har vi på de fleste områder ikke funnet signifikante forskjeller i auskultasjonsferdigheter mellom to grupper, leger og studenter, med betydelig forskjell i klinisk erfaringsnivå. Spesielt er positiv prediktiv verdi ikke signifikant bedre hos leger sammenliknet med studenter. Da er også pediatere med mer enn 25 års erfaring inkludert i legegruppen. Det vil bety at erfarne barneleger ville henvist like mange barn uten organisk hjertefeil som leger med liten erfaring basert på vurdering av hjertelydene denne studien har presentert.

Den positive prediktive verdien for identifisering av patologisk hjertebilyd for leger og studenter er henholdsvis 0,32 og 0,29. Det vil si at dersom de representerer førstelinjetjenesten vil ca. 30 % av barna de velger å henvise ha hjertefeil. Dette er bedre enn det som rapporteres fra de barnekardiologiske poliklinikkene hvor forekomsten av organisk hjertefeil blant de henviste er mellom 3% (2) og 10% (1). En viktig grunn til at den positive prediktive verdi er høyere i vår studie enn det som observeres ved henvisinger er at andelen av patologi i vårt materiale er vesentlig høyere enn det en allmennpraktiker eller helsestasjonslege møter i sin praksis. Økt prevalens av den tilstand som skal identifiseres vil nødvendigvis øke positiv prediktiv verdi. Man må også ta høyde for at det i klinisk praksis ikke bare er mistanke om organisk hjertefeil som gjør at man velger å henvise barn med bilyd. Spesielt kan en henvisning for endelig avklaring med ekkokardiografi være med på å berolige engstelige foreldre.

Vi har ikke funnet signifikante forskjeller mellom leger med mindre enn 7 års erfaring og studenter når det gjelder treffsikkerhet eller positiv prediktiv verdi. Dette gjelder også sensitivitet på patologiske lyder. Dermed ville ikke studentene hverken oversett flere patologiske eller henvist flere friske enn leger med opptil 7 års erfaring. Dette stemmer med funn i andre studier (14), og støtter vår bruk av medisinstudenter som representanter for førstelinjetjenesten. Derimot fant vi en signifikant økning i positiv prediktiv verdi når vi testet to grupper av leger med henholdsvis mer eller mindre enn 7 års erfaring. Positiv prediktiv verdi økte da fra 0,28 i gruppen med mindre enn 7 års erfaring til 0,36 i gruppen med mer enn 7 års erfaring. Det kan imidlertid diskuteres om denne forskjellen er klinisk signifikant, da de fleste identifiserte som syke allikevel vil være hjertefriske. Det er signifikante forskjeller mellom legene og studentene når det kommer til sensitivitet på vurdering av patologisk bilyd. Det vil si at legene ville gått glipp av færre tilfeller med organisk hjertefeil enn studentene. Imidlertid

forsvinner denne forskjellen når de mest erfarne legene tas ut av legegruppen. Det ser derfor ut til at det må lang erfaring til før man blir signifikant bedre enn studenter i å selektere ut pasienter for henvisning. Selv denne erfarne gruppen henviser flere barn med fysiologisk bilyd enn barn med faktisk hjertepatologi, og overser mer enn 20% av barna med patologisk hjertebilyd.

Det er en signifikant forskjell mellom leger og studenter på spesifisitet for identifisering av normal hjertelyd. Dette kan også sees ut fra figur 6, hvor legene kategoriserer et svært lite antall av bilydene som normale. Legene er derfor noe dyktigere på å skille bilyd fra ikke bilyd. Men da det store flertall av bilyder er fysiologiske vil ikke dette være like viktig som å kunne skille fysiologisk fra patologisk bilyd og her er legene ikke tydelig bedre enn studentene.

Både leger og studenter mener å høre bilyd hos en tredel av klinisk friske barn med normal hjertelyd. Hos både leger og studenter finner vi at ca 70% av de henviste barna henvises uten at det foreligger hjertefeil. Vi finner lik treffsikkerhet i alle gruppene på ca 60%. Imidlertid er ikke treffsikkerhet alene et godt mål. Den totale andelen av normale hjertelyder i vårt materiale er på 68%. Det vil si at dersom man ikke tar stilling til en hjertelyd og klassifiserer alle lyder som normale, vil man få en treffsikkerhet som er høyere enn om man hadde tatt stilling til lydene. Dette gjelder for alle undersøkelser hvor forekomsten av det man undersøker er lav. Da fanger man imidlertid ikke opp de som har hjertefeil, men som ved undersøkelsestidspunktet ikke har andre kliniske tegn på hjertefeil enn bilyd. Dermed må man mulighens akseptere at det henvises mange friske, dersom man skal plukke opp flest med organisk hjertefeil.

Interobservatørkappa var 0,32 for legene og 0,18 for studentene. Flere studier har funnet inter- og intraobservatørkappaverdier på over 0,8 (2,12) når det gjelder klassifiseringen av hjertelyder. Samtidig er det publisert en studie hvor interobservatørkappa var på bare 0,05 (11). I lys av dette er ikke nødvendigvis våre verdier spesielt lave. Felles for vår studie og studien som fant stor interobservatørvariasjon er det store antallet testpersoner. Dette var også en langt mer homogen gruppe enn vår, med testing av 30 privatpraktiserende pediatere som alle var ferdige spesialister. Vår lave kappaverdi kn også forklares utfra lydmaerialet, som er spesifikt valgt ut med tanke på at de skal være vanskelige å skille.

Positiv prediktiv verdi økte fra 0,24 i seanse A til 0,33 i seanse B for studentene. Det samme gjaldt legene, hvor positiv prediktiv verdi økte fra 0,29 i seanse A til 0,36 i seanse B. Den mest sannsynlige årsaken til denne økningen er den lavere hyppigheten av fysiologisk bilyd i seanse

B. Da fysiologisk bilyd er vanskelig å vurdere vil det ved en lavere forekomst være færre forstyrrende elementer, fysiologisk bilyd, å forholde seg til. Dette spesielt fordi vårt utvalg av patologiske hjertelyder ligger innenfor karaktertrekket til de fysiologiske bilydene.

Vi var også interessert i å finne ut om rekkefølgen av lydene kunne ha betydning for besvarelsene. Da de 11 siste lydene i seanse B er normale, kan man stille spørsmål om selve rekkefølgen på lydene i testen har betydning for testdeltakerenes vurdering, såkalt tidsbias. Av figur 6 sees at samsvaret mellom de like lydene i seanse A og B ikke endres spesielt i løpet av seanse B. Vi ser også at kurvene for leger og studenter følger hverandre gjennom seansen. Det vil si at der hvor samsvaret var høyt for legene, var den også det for studentene. Vi finner derfor ikke et tidsbias vår undersøkelse. To lydpar bør imidlertid kommenteres nærmere: 1) Hos legene sees et fall i samsvar på lydpar a2n/b6n, en normal lyd, som 52% klassifiserte korrekt i seanse A og 12% klassifiserte korrekt i seanse B. Med bakgrunn i det lave samsvaret på dette lydparet, gikk vi etter avsluttet testing gjennom dette lydparet på nytt for å verifisere at lydene faktisk er identiske. 2) Det høyeste samsvaret sees mellom a21p og b13p, en ASD grad III, som er den kraftigste bilyden i materialet. Dette er også den enkelte patologiske bilyd som flest deltakere identifiserer korrekt. Den skiller seg således ut på styrke i forhold til de andre bilydene i materialet. Dermed kan styrken alene ha vært avgjørende for vurderingen, ettersom fysiologiske bilyder sjeldent er så sterke.

Svakheter

Den største svakheten vi kan se i vår studie er at det ikke er gjort ekkokardiografisk undersøkelse på de barna som ut fra gullstandarden fikk diagnostisert en fysiologisk bilyd. Da kan man stille spørsmålsteget ved om gullstandarden er riktig. Det er spesielt én lyd hvor både studenter og leger er uenig i gullstandard. Det er lyd a24n som kun 20% av studentene og ingen av legene bedømte korrekt. Det kan derfor være noe usikkerhet rundt gullstandarden på denne lyden. Vi har likevel besluttet å beholde lyden i materialet, da det store antallet lyder gjør at bedømmelsen av lyden ikke vil påvirke resultatet i stor grad.

Det bør nevnes at en testsituasjon ikke vil være identisk med en undersøkelse hvor legen har barnet foran seg. Det er naturlig å tenke at i en slik situasjon vil positiv prediktiv verdi bli bedre. På den annen side er vår studie reell for de mottakende lege i en telemedisinsk setting. Da vil mottakende lege kun ha lyden å forholde seg til.

I vår studie er forekomsten av patologisk hjertelyd på 10%. Dette er vesentlig høyere enn hva som er reell forekomst av patologi for henvisende lege på en helsestasjon. Ved bruk av telemedisinsk henvisning vil imidlertid forekomsten av patologiske bilyder kunne være lik den forekomsten som den vurderende lege vil møte i sitt arbeide (1).

Det vil kunne bli hevdet at bruken av studenter som testpersoner ikke gir et riktig bilde av primærlegers auskultasjonsferdigheter. Denne bruken kan forsvares i studier (14) og ut fra våre resultater. Det er etter statistiske beregninger få og små forskjeller mellom leger og studenter.

Det kan argumenteres for at det er barnekardiologer som vil sitte som mottagende lege, og at disse vil prestere bedre. Disse må i så fall prestere betraktelig bedre enn erfarne barneleger, dersom telemedisinsk vurdering skal være praktisk mulig. Det er vist i studier at barnekardiologer har bedre auskultasjonsferdigheter enn det vi finner, men disse er testet på en populasjon med høy andel grov patologi og er dermed ikke representativ for den populasjon legene vil møte i sitt arbeide angående telemedisin. Det er i all hovedsak vanskelige tilfeller, hvor legen er usikker på om den hørte bilyden er patologisk eller ikke, at en telemedisinsk vurdering er aktuell. Derfor er det ikke interessant å bruke studier hvor mottagende lege har vurdert tilsendte hjertelyder inneholdende mye grov patologi. Dersom man skal bruke spesialister i barnekardiologi til å vurdere hjertelyder telemedisinsk, må deres ferdigheter kartlegges nærmere. Det vil være interessant å teste barnekardiologer ved universitetssykehusene på et lydsett som likner vårt.

Konsekvenser

Det er i vår studie ikke funnet noen tydelig forskjell mellom leger og studenter når det gjelder presisjon i utvelgelse av barn med hjertefeil basert på vurdering av testlyder tatt opp på klinisk friske barn på en helsestasjon. Vår studie forteller oss at både leger og studenter er rimelig gode til å skille mellom bilyd og ikke-bilyd. Ingen av gruppene er imidlertid spesielt gode til å skille mellom fysiologisk bilyd og patologisk hjertelyd når punctum maximum og styrke er nesten identisk slik de var i vårt testmateriale av hjertelyder. Dette sier oss to ting: 1) Det er forståelig at mange barn henvises uten funn til utredning for hjertebilyd. 2) Det er fornuftig at mottakende leger utfører ekko-dopler undersøkelse av alle disse barna, da lydvurdering alene er for usikkert.

Ved sammenlikning av to legegrupper med forskjellig erfaringsnivå finner vi at økning i erfaring bedrer presisjonen for identifikasjon av patologi, men ikke opp mot en presisjon som

er klinisk anvendelig. Ved bruk av telemedisinsk vurdering vil det være av største betydning at mottakende lege er vesentlig dyktigere enn henvisende lege. At mottakende lege faktisk er bedre enn henvisende lege må kvalitetssikres da barneleger, selv med erfaring på over 20 år, ikke er vesentlig bedre enn mindre erfarne leger. Telemedisinsk vurdering av hjertelyder synes ikke egnet for de situasjoner hvor det foreligger differensialdiagnostikk mot fysiologisk bilyd. Da må hele barnet sendes til vurdering, ikke bare hjertelyden.

Økt trening av leger til å karakterisere hjertelyder synes ikke å øke presisjonen i en slik grad at telemedisinsk vurdering av lyder blir praktisk mulig. Mottakende lege må fortsatt foreta ekkokardiografisk undersøkelse.

En annen mulighet kan vise seg å være bruken av såkalte kunstige nevrale nettverk (Artificial Neural Network, ANN). ANN baserer seg på at datamaskiner kan læres opp til å gjenkjenne mønstre (pattern recognition) i lydene, i motsetning til en kognitiv vurdering. Dersom disse har en stor nok database av lyder som er korrekt knyttet opp mot en fasit, vil disse kunne bestitte et langt større erfaringsgrunnlag enn det en lege i førstelinjetjenesten kan oppnå. Studier har vist at disse nevrale nettverkene med sensitivitet og spesifisitet på opptil 100% kan skille fysiologisk fra patologisk bilyd (19). Dette vil dessuten langt raskere enn å sende lyden.

Konklusjon

Våre funn viser at det er vanskelig å skille svake bilyder. Derfor er det forståelig at mange barn henvises uten funn til utredning for hjertebilyd. Telemedisinsk vurdering av hjertelyder synes ikke egnet, da mottakende legers ferdigheter ikke er spesielt bedre enn den henvisendes. Da vil det også være fornuftig at mottakende leger utfører ekko-dopler undersøkelse av alle disse barna, da lydvurdering alene er for usikkert.

Hovedfunn

- Vår studie kan ikke påvise tydelige forskjeller i auskultasjonsferdigheter mellom medisinstudenter og leger.
- Vi har funnet en positiv prediktiv verdi som er høyere enn det man ser i praksis,
- Det er sammenheng mellom legens erfaring og vedkommendes positive prediktive verdi, men forskjellen er ikke så stor at den burde få praktiske konsekvenser.
- Prevalens av fysiologisk bilyd påvirker positiv prediktiv verdi.

Hovedbudskap

- Det er vanskelig å plukke ut patologiske bilyder. Dermed er det forståelig at mange av barna som henvises til barnekardiolog, ikke har funn når de undersøkes av mottakende lege.
- Det er forståelig at mottakende lege utfører ekko-dopler undersøkelse av alle henviste barn, da lydvurdering alene er for usikkert.
- Telemedisinsk vurdering av hjertelyder synes ikke egnet for de situasjoner hvor det foreligger differensialdiagnostikk mot fysiologisk bilyd.
- Dersom man skal bruke spesialister i barnekardiologi til å vurdere hjertelyder telemedisinsk, må deres ferdigheter kartlegges nærmere.
- Evt. kan man i fremtiden vurdere å ta i bruk andre metoder for å hjelpe henvisende lege til bedre presisjon på henvisingene (eks ANN).

Referanser:

- 1) Norgard G, Greve G, Rosland GA, Berg A. Henvisningspraksis og klinisk vurdering av bilyder hos barn. Tidsskr Nor Laegeforen. 2005;125:996-998.
- 2) Dahl LB, Hasvold P, Arild E, Hasvold T. Heart murmurs recorded by a sensor based electronic stethoscope and e-mailed for remote assessment. Arch Dis Child. 2002;87:297-301; discussion 297-301.
- 3) McIntosh N, Helms PJ, Smyth R. Forfar and Arneil's Textbook of Pediatrics. Churchill Livingstone; 2003. p. 836.
- 4) Kliegman RM, Behrman RE, Jenson HB, Stanton BF. Nelson Textbook of Pediatrics. Saunders; 2007. p. 1878.
- 5) Meberg A, Hals J, Thaulow E. Congenital heart defects--chromosomal anomalies, syndromes and extracardiac malformations. Acta Paediatr. 2007;96:1142-1145.
- 6) Kliegman RM, Behrman RE, Jenson HB, Stanton BF. Nelson Textbook of Pediatrics. Saunders; 2007. p. 1864.
- 7) Wilkinson J. Assessment of the infant and child with suspected heart disease. In: Robinson MJ, Robertson DM (eds). Practical Paediatrics 6th ed. Churchill Livingstone. Edinburgh 2007: 517-25.
- 8) Berg A, Greve G, Hirth A, Rosland GA, Norgard G. Vurdering av barn med hjertebilyder. Tidsskr Nor Laegeforen. 2005;125:1000-1003.
- 9) Mahnke CB, Nowalk A, Hofkosh D, Zuberbuhler JR, Law YM. Comparison of two educational interventions on pediatric resident auscultation skills. Pediatrics. 2004;113:1331-1335.
- 10) Gaskin PR, Owens SE, Talner NS, Sanders SP, Li JS. Clinical auscultation skills in pediatric residents. Pediatrics. 2000;105:1184-1187.
- 11) Haney I, Ipp M, Feldman W, McCrindle BW. Accuracy of clinical assessment of heart murmurs by office based (general practice) paediatricians. Arch Dis Child. 1999;81:409-412.
- 12) Rajakumar K, Weisse M, Rosas A et al. Comparative study of clinical evaluation of heart murmurs by general pediatricians and pediatric cardiologists. Clin Pediatr (Phila). 1999;38:511-518.
- 13) Barrett MJ, Kuzma MA, Seto TC et al. The power of repetition in mastering cardiac auscultation. Am J Med. 2006;119:73-75.
- 14) Vukanovic-Criley JM, Criley S, Warde CM et al. Competency in cardiac examination skills in medical students, trainees, physicians, and faculty: a multicenter study. Arch Intern Med. 2006;166:610-616.
- 15) McCrindle BW, Shaffer KM, Kan JS, Zahka KG, Rowe SA, Kidd L. An evaluation of parental concerns and misperceptions about heart murmurs. Clin Pediatr (Phila). 1995;34:25-31.
- 16) Holmstrøm H, Thaulow E. Ekko for alle barn med bilyd? Tidsskr Nor Laegeforen. 2005;125:994.
- 17) Fleiss JL, Levin B, Paik MC, Fleiss J. Statistical Methods for Rates & Proportions. The measurement of inter-rater agreement. Wiley-Interscience; 2003. p. 610-617.
- 18) Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.
- 19) DeGroff CG, Bhatikar S, Hertzberg J, Shandas R, Valdes-Cruz L, Mahajan RL. Artificial neural network-based method of screening heart murmurs in children. Circulation. 2001;103:2711-2716.

Vedlegg 1: Informasjon til deltakerne

Identifisering av hjertebilyder hos barn, seks uker til to år gamle, tatt opp på helsestasjonen.

Bakgrunn

På helsestasjon blir barn undersøkt for bilyd. Fysiologisk bilyd forekommer relativt ofte mens patologisk bilyd opptrer sjeldent. Det er en utfordring for helsestasjonslege å avgjøre om en bilyd er fysiologisk eller patologisk. Det kan også være vanskelig å avgjøre med sikkerhet om det forekommer bilyd eller ikke.

Din vurdering av 2x25 hjertelyder tatt opp på helsestasjon for barn

Vi ber om at du vurderer 25 hjertelyder tatt opp på barn i alderen seks uker til to år. Alle barna er klinisk friske. Hjertelyden er tatt opp i området mellom venstre mamma og venstre sternalrand (4de interkostalrom venstre sternalrand) eller ved punktum maksimum ved bilyd.

Metode

Først ber vi om å sikre deg at du hører lydene godt med gode øretelefoner. Bilyder har lav frekvens (bass lyd). Derfor må du sjekke at du kan høre referanselydene på 30, 150 og 800 Hz godt.

Deretter ber vi om at du hører på lydene enkeltvis og avgjør om lyden er:
Normal hjertelyd, Fysiologisk bilyd, Patologisk bilyd eller om du ikke kan ta stilling til lyden.

På vedlagte skjema er det en tabell hvor du fører opp lydnummeret i den rubrikk som svarer til din vurdering.

Det er to grupper av hjertelyder, hhv seanse A og B hvor bilydene forekommer med forskjellig hyppighet. Vi ber deg vurdere lydene fra begge seansene.

To leger har før hørt på hjertelydene, den ene er helsestasjonslege, den andre er barnekardiolog, begge er spesialister i barnesykdommer. I henhold til deres vurdering kan din vurdering klassifiseres som rett eller gal og din treffsikkerhet kan beregnes. Dette gjøres etter at hele undersøkelsen er utført.

Du kan avtale med den som administrerer undersøkelsen hvordan du kan få tilbakemelding på din treffsikkerhet i forhold til gjennomsnittet for treffsikkerhet i denne undersøkelsen.

Vi vil takke deg for din delaktighet.

Ditt bidrag kan ha betydning for bedring av henvisningsrutiner for barn med bilyd.

Vedlegg 2: Svarskjema leger

Kjønn(K/M) ____ Ferdig lege i år: ____ Ferdig spesialist i pediatri ____
 Tilknyttet hjerteseksjonen ____ Fordypningsstilling: ____

SEANSE A

Bedømt som	Normal hjertelyd		Fysiologisk bilyd		Patologisk bilyd	
Nummer på hjertelyden som vurderes						
Antall vurderinger						
Vurdering mot en standard	Rett	Gal	Rett	Gal	Rett	Gal

SEANSE B

Bedømt som	Normal hjertelyd		Fysiologisk bilyd		Patologisk bilyd	
Nummer på hjertelyden som vurderes						
Antall vurderinger						
Vurdering mot en standard	Rett	Gal	Rett	Gal	Rett	Gal

Prosjektansvarlig: Per Lagerløv, førsteamanuensis i allmennmedisin, Universitetet i Oslo
 Prosjektmedarbeider: Leif Brunvand, avdelingsoverlege, Kvinne-Barn Sentret Ullevål Universitetssykehus
 Stud med. Håkon B Jegstad og Geir Kristensen, Universitetet i Oslo

Vedlegg 3: Svarskjema studenter

Kjønn(K/M) ____ Kull: _____ Alder: _____

SEANSE A

Bedømt som	Normal hjertelyd		Fysiologisk bilyd		Patologisk bilyd	
Nummer på hjertelyden som vurderes						
Antall vurderinger						
Vurdering mot en standard	Rett	Gal	Rett	Gal	Rett	Gal

SEANSE B

Bedømt som	Normal hjertelyd		Fysiologisk bilyd		Patologisk bilyd	
Nummer på hjertelyden som vurderes						
Antall vurderinger						
Vurdering mot en standard	Rett	Gal	Rett	Gal	Rett	Gal

Prosjektansvarlig: Per Lagerløv, førsteamanuensis i allmennmedisin, Universitetet i Oslo

Prosjektmedarbeider: Leif Brunvand, avdelingsoverlege, Kvinne-Barn Sentret Ullevål Universitetssykehus

Stud med. Håkon B Jegstad og Geir Kristensen, Universitetet i Oslo